

PTO: 2003-5030

Japanese Published Unexamined Patent Application (A) No. 06-141309, published May 20, 1994; Application Filing No. 04-290019, filed October 28, 1992; Inventor(s): Atsuyuki Mizuno; Assignee: Canon Corporation; Japanese Title: Image/Voice Communication Terminal Devices

IMAGE/VOICE COMMUNICATION TERMINAL DEVICES

CLAIM(S)

1) An image/voice communication terminal device equipped with a direction-changeable image display unit, and comprising: a voice sample collecting means for collecting the voice samples at multiple positions around the image/voice communication terminal device; a target direction-determining means for determining the target direction by comparing the sound volumes of voice samples collected by said voice sample collection means; a display direction-changing means for changing the display direction of said display means to the target direction determined by said target direction-determining means.

2) An image/voice communication terminal device equipped with a photographing direction-changeable video camera, and comprising: a voice sample collecting means for collecting the voice samples at multiple positions around the image/voice communication terminal device; a target

direction-determining means for determining the target direction by comparing the sound volumes of voice samples collected by said voice sample collection means; a photographing direction-changing means for changing the photographing direction of said video camera to the target direction determined by said target direction-determining means.

3) An image/voice communication terminal device equipped with an image display direction-changeable means and a photographing direction-changeable video camera, and comprising: a voice sample collecting means for collecting the voice samples at multiple positions around the image/voice communication terminal device; a target direction-determining means for determining the target direction by comparing the sound volumes of voice samples collected by said voice sample collection means; a direction changing means for changing the display direction of said display means and the photographing direction of said video camera to the target direction determined by said target direction-determining means.

DETAILED DESCRIPTION OF THE INVENTION

(0001)

(Field of Industrial Application)

The present invention pertains to an image/voice terminal device.

(0002)

(Prior Art)

Due to development of a communication network such as that represented by an ISDN, a quantity of data that can be communicated per a time unit is increasing and terminals that conduct communications by using animate images and voices are emerging.

(0003)

The terminal of this type is generally equipped with a display unit for displaying the reception image data, such as a CRT and a LCD, and is generally equipped with a video camera as a means of collecting image data for transmitting.

(0004)

(Problems of the Prior Art to Be Addressed)

The aforementioned display unit, however, has a limited visible range, particularly, a limited visible angle, so the user has to conduct the communications within the visible range of display unit. If the user moves

away during the communication, the display unit goes out of his/her sight, which is a problem.

(0005)

When the video is fixed in position, the user has to move to the photographing range of the video camera. To solve this problem, it was considered to make the video camera-mounting unit movable. In such a case, however, it will be tedious if the user has to change the direction.

(0006)

In addition, because of these problems, the user has to be too conscious of the terminal, which makes the communication environment uncomfortable and unnatural.

(0007)

The present invention, to solve the aforementioned problems, attempts to present an image/voice communication terminal device that the user can use naturally without being too conscious of the visible range of a display means and the photographable range of a video camera.

(0008)

(Means to Solve the Problems)

To accomplish the aforementioned objective, the present invention structures the image/voice communication terminal device as described in the following 1), 2), 3).

(0009)

1) The image/voice communication terminal device equipped with the direction-changeable image display unit comprises: a voice sample collecting means for collecting the voice samples at multiple positions around the image/voice communication terminal device; a target direction-determining means for determining the target direction by comparing the sound volumes of the voice samples collected by said voice sample collection means; a display direction-changing means for changing the display direction of said display means to the target direction determined by said target direction-determining means.

(0010)

2) The image/voice communication terminal device equipped with the photographing direction-changeable video camera comprises: a voice sample collecting means for collecting the voice samples at multiple positions around the image/voice communication terminal device; a target

direction-determining means for determining the target direction by comparing the sound volumes of the voice samples collected by said voice sample collection means; a photographing direction-changing means for changing the photographing direction of said video camera to the target direction determined by said target direction-determining means.

(0011)

3) The image/voice communication terminal device equipped with the display direction-changeable image display means and with the photographing direction-changeable video camera comprises: a voice sample collecting means for collecting the voice samples at multiple positions around the image/voice communication terminal device; a target direction-determining means for determining the target direction by comparing the sound volumes of the voice samples collected by said voice sample collection means; a direction changing means for changing the display direction of said display means and the photographing direction of said video camera to the target direction determined by said target direction-determining means.

(0012)

(Operation)

According to the aforementioned structure (1), the direction of the user is determined based on the sound volume of the voice samples collected at multiple positions around the terminal, and the display direction of the display means is directed toward said direction. According to the structure (2), the direction of the user is determined based on the sound volumes of the voice samples collected at multiple positions around the terminal. The photographing direction of the video camera is directed to the target direction. According to the structure (3), the display means and the video camera are directed to the user by the result of determination output in (1) and (2).

(0013)

(Embodiment Example)

The embodiment example of the present invention is explained below in detail.

(0014)

(Example 1)

Fig. 1 shows an oblique view of the contour of the “image/voice communication terminal device” as one embodiment example of the present

invention. In the figure, 1 and 2 show the microphones, respectively, and 8 the movable display unit.

(0015)

Fig. 2 shows a block diagram of the structure of the embodiment example. In the figure, 1 and 2 show the microphones, respectively, which are the voice input means, and 3 a speaker, which is a voice output means.

(0016)

In the figure, 4 indicates a voice interface section for performing a switching operation, by the command from the system controller 14, between microphones 1 and 2 and speaker 3 as the voice input/output means, an echo canceling operation for deleting the echo when the microphones 1 and 2 and speaker 3 are used, a transfer operation for transferring the voice data collected from the microphones 1 and 2 as the voice sample collection means to the system controller 14, and an operation of generating a dial tone, call sound, busy tone, and arrival sound tone.

(0017)

In the figure, 5 indicates a voice encoding/decoding section for encoding the transmission voice signal and decoding the reception voice signal by the command from the controller 14, according to the voice signal-encoding/decoding algorithm of 64 kbpsPCM A-law, 64 kbps PCM μ -law,

64 bps/56 kbps/48 kbps SB-ADPCM, 32 kbps ADPCM, 16 kbps (for example, APC – AB), and 8 kbps.

(0018)

In the figure, 6 indicates a video camera, which is an image input means, for inputting a self-image, 7 an image/character camera, which is a image input means, for inputting pictures and graphs, and 8 a movable display unit for displaying the input image from the camera 6 or image/character camera 7, and the received image, and the operation screen.

(0019)

In the figure, 9 indicates the video interface section for performing, by the command from the system controller 14, an image input means-switching operation, a display switching operation of input image, received image, and of operation screen display, an image signal synthesizing operation for dividedly displaying them on the display unit 8, and a camera control operation for controlling the display direction of the display unit 8.

(0020)

In the figure, 10 indicates the video encoding/decoding section for encoding the transmission image signal and decoding the received image signal according to the CCITT proposal H.261, and 11 the data terminal for performing the data communication. And 12 indicates the data interface

section that provides multiplexing/separating section 15 of the transmission data from the data terminal 11 and system controller 14, and provides the received data to the data terminal 11 or system controller 14.

(0021)

In the figure, 13 indicates an operation unit, such as a keyboard or a touch panel, for inputting the control data for controlling the entire device; 14 indicates the system controller which is equipped with a CPU, ROM, RAM, and with an auxiliary memory device, and which creates the operation/display screen according to the control status of the entire device by monitoring each section, executes the application program, commands the voice sample collection to the voice interface section 4, compares the sound volume data of the voice sample transmitted from the voice interface section 4, and teaches the target direction determined by the comparison result to the video interface section 9.

(0022)

In the figure, 15 indicates the multiplexing/separating section that multiplexes the voice signal from the voice encoding/decoding section 5, image signal from the video signal encoding/decoding section 10, the data from the data interface section 12, and the BAS from the system controller 14 into the transmission frame unit, and that separates them into each

medium in a reception frame unit to inform them to each section, according to the CCITT proposal H. 221. In the same figure, 16 indicates the circuit interface section for controlling the circuit according to the ISDN user network interface.

(0023)

Fig. 3 indicates a flowchart of the operation of the embodiment example. In the figure, the steps in S2 – S6 are called “an automatic display section-tracking operation.” The automatic display section-tracking operation is automatically activated at the start of the communication and can be activated during the communication whenever the operation unit 13 commands it.

(0024)

First, the circuit interface section 16 controls the transmission and reception and, thus, the communication starts (S1). Then, the system controller 14 commands, via the voice interface section 4, the voice sample collection to the microphones 1 and 2. The microphones 1 and 2 begin to collect the voice samples (S2). The collected voice samples are transmitted to the system controller 14 via the voice interface section 4. The system controller 14 compares the voice samples transmitted from the microphones 1 and 2, respectively, and determines which sound volume is higher (S3).

Based on the result of comparison, the system controller 14 determines the direction of the target, i.e., the speaker (S4), and commands the video interface section 9 to direct the movable display unit 8 to the target (S5). The video interface section 9 directs the movable display unit 8 to the commanded direction (S6).

(0025)

Hereinafter, the automatic display unit-tracking operation S2 – S6 is repeated until the communication is completed (Yes in S8) if the operation section 13 commands the automatic display unit-tracking operation.

(0026)

(Embodiment Example 2)

In the aforementioned embodiment example 1, two microphones were used for the voice sample collecting means.

(0027)

By contrast, in this example, one movable microphone is used for the voice sample collecting means, and when the voice samples are collected, the system controller 14 moves/controls the movable microphone to collect the voice samples at several positions.

(0028)

(Embodiment Example 3)

Fig. 4 shows an oblique view of the “image/voice communication terminal device as embodiment example 3. In the figure, 1 and 2 indicate the microphones, and 6 – 1 the movable video camera.

(0029)

Fig. 5 shows a block diagram of the structure of the embodiment example 3. In the figure, blocks other than 6 – 1, 8 – 1, 9 – 1, 14 – 1, indicate the same blocks as those in Fig. 1, so they are omitted from the explanation.

(0030)

In the figure, 6- 1 indicates an image input means and a movable video camera for inputting a self-image; 8-1 indicates the display unit for displaying the image input from the camera 6-1 or picture image camera 7, received image from the other party, and operation screen.

(0031)

In the figure, 9 – 1 indicates the video interface section that performs, by the command from the system controller 14 – 1, the switching operation of image input means, the switching operation of input images, the reception images, and of operation screen displays, the image signal

synthesizing operation for dividedly displaying them on the display section 8 – 1, and the camera control operation for controlling the direction of movable video camera 6.

(0032)

In the figure, 14 – 1 indicates the system controller which is equipped with a CPU, ROM, RAM, and with an auxiliary memory device, and which creates the operation/display screen according to the control status of the entire device by monitoring the status of each section, executes the application program, command the voice sample collection to the voice interface section 4, compares the voice data of the voice samples transmitted from the voice interface section, and provides the video interface section 9 – 1 with the direction of the target determined by the result of the comparison.

(0033)

Fig. 3 shows a flowchart indicating the operation of the embodiment example. In the figure, the operations in the steps S12 – S16 are put together and called “an automatic camera-tracking operation.” The automatic camera-tracking operation is automatically activated at a time of starting the communication, and can be activated during the communication whenever the operation unit 13 commands it.

(0034)

First, the circuit interface section 16 controls the transmission or reception and, thus, the communication starts (S11). Then, the system controller 14 – 1 commands, via the voice interface section 4, the voice sample collection to the microphones 1 and 2. Here, the microphones 1 and 2 begin to collect the voice samples around the terminal (S12). The collected voice samples are transferred to the system controller 14 – 1 via the voice interface section 4. The system controller 14 – 1 compares the voice samples transmitted from the microphones 1 and 2, respectively to determine which sound volume is higher (S13). Based on the result of comparison, the system controller 14 – 1 determines the direction of the target, i. e., the speaker (S14), and commands the video interface section 9 – 1 to direct the movable video camera 6-1 to the direction of the target (S15). The video interface section 0-1 directs the movable video camera 6-1 to the commanded direction.

(0035)

Hereinafter, the automatic camera-tracking operations S12 – S16 are repeated until the communication ends if the operation section 13 commands the automatic camera-tracking operation (Yes in S17).

(0036)

(Embodiment Example 4)

In the aforementioned Embodiment Example 3, two microphones were used for the voice sample collection means.

(0037)

In this example, one movable microphone is used for the voice sample collection means. When the voice samples are collected, the system controller 14 – 1 moves/controls the movable microphone to collect the voice sample at several positions.

(0038)

In the aforementioned each embodiment example, the display unit or video camera is directed to the user, but the present invention is not limited to this method; the display unit and video camera can both be directed to the user. Also, the direction can be changed by rotating the mounting unit, or by changing the reflection direction by using a mirror.

(0039)

(Advantage)

As explained above, according to the present invention, the image display means and video camera automatically track down the user, the user needs not be conscious of the visible range of the display means or the

photographing range of the video camera, which allows comfortable communications.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows an oblique view of the contour of the embodiment example 1. Fig. 2 shows a block diagram of the embodiment example 1. Fig. 3 shows a flowchart of the embodiment example 1. Fig. 4 shows an oblique view of the embodiment example 3. Fig. 5 shows a block diagram of the embodiment example 3. Fig. 6 shows a flowchart of the embodiment example 3.

1, 2. microphones

8. movable display unit

9. video interface section

14. system controller

Translations
U. S. Patent and Trademark Office
8/19/03
Akiko Smith

(19)日本国特許庁(JP)

(12) 公開特許公報(A)

(11)特許出願公開番号

特開平6-141309

(43)公開日 平成6年(1994)5月20日

(51)Int.Cl. ⁵	識別記号	庁内整理番号	FI	技術表示箇所
H 0 4 N	7/14	8943-5C		
	5/222	Z		
	5/225	Z		

審査請求 未請求 請求項の数3(全 9 頁)

(21)出願番号 特願平4-290019

(22)出願日 平成4年(1992)10月28日

(71)出願人 000001007

キャノン株式会社

東京都大田区下丸子3丁目30番2号

(72)発明者 水野 敦之

東京都大田区下丸子3丁目30番2号 キャ
ノン株式会社内

(74)代理人 弁理士 丹羽 宏之 (外1名)

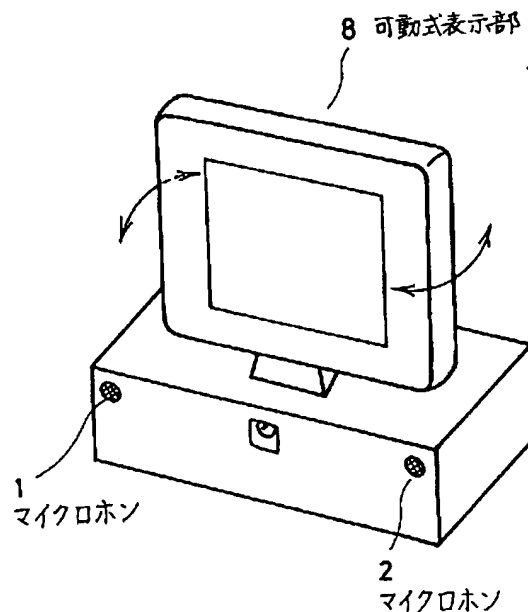
(54)【発明の名称】 画像・音声通信端末装置

(57)【要約】

【目的】 使用者が表示手段の視認可能範囲或はビデオカメラの撮像可能範囲を特に意識することなく、自然に端末を使用することができる画像・音声通信端末装置を提供する。

【構成】 画像・音声通信端末装置の周囲の複数位置における音声サンプルをマイクロホン1, 2で拾集し、この音声サンプルの音量を比較して使用者の方向を判定し、その方向に可動式表示部8を向ける。この自動追尾により前記目的が達成できる。

実施例1の外観斜視図



【特許請求の範囲】

【請求項1】 表示方向を変更可能な画像等の表示手段を備えた画像・音声通信端末装置であって、この画像・音声通信端末装置の周囲の複数位置において音声サンプルを拾集する音声サンプル拾集手段と、この音声サンプル拾集手段で拾集した音声サンプルの音量を比較して目標方向を判定する目標方向判定手段と、この目標方向判定手段で判定した目標方向へ前記表示手段の表示方向を変更させる表示方向変更手段とを備えたことを特徴とする画像・音声通信端末装置。

【請求項2】 撮像方向を変更可能なビデオカメラを備えた画像・音声通信端末装置であって、この画像・音声通信端末装置の周囲の複数位置において音声サンプルを拾集する音声サンプル拾集手段と、この音声サンプル拾集手段で拾集した音声サンプルの音量を比較して目標方向を判定する目標方向判定手段と、この目標方向判定手段で判定した目標方向へ前記ビデオカメラの撮像方向を変更させる撮像方向変更手段とを備えたことを特徴とする画像・音声通信端末装置。

【請求項3】 表示方向を変更可能な画像等の表示手段と、撮像方向を変更可能なビデオカメラを備えた画像・音声通信端末装置であって、この画像・音声通信端末装置の周囲の複数位置において音声サンプルを拾集する音声サンプル拾集手段と、この音声サンプル拾集手段で拾集した音声サンプルの音量を比較して目標方向を判定する目標方向判定手段と、この目標方向判定手段で判定した目標方向へ、前記表示手段の表示方向と前記ビデオカメラの撮像方向を変更させる方向変更手段とを備えたことを特徴とする画像・音声通信端末装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、画像・音声通信端末装置に関する。

【0002】

【従来の技術】近來ISDNに代表されるような通信網の発達によって、単位時間当たりに通信できるデータ量が急激に増えつつあり、動画像と音声を用いて通信を行うような端末が現れている。

【0003】このような端末では、一般にCRTやLCDといった受信画像データ等の表示部を備えており、また送信する画像データを拾集する手段として、一般にビデオカメラを備えている。

【0004】

【発明が解決しようとする課題】しかし前述のような表示部は、視認可能範囲特に角度が限られているため、使用者は表示部が見える範囲で通信や操作を行わなければならない。あるいは通信中に使用者が移動した場合などには表示部が見えなくなってしまうといった問題がある。

【0005】また、ビデオカメラが固定されている場合

は、使用者がビデオカメラの撮像可能範囲内に移動しなければならない。あるいは、この問題を一部解決するものとしてビデオカメラの取付部を可動式にすることも考えられるが、この場合には、使用者がその方向を変更するのに手間がかかるという問題がある。

【0006】さらに、このような問題のため、使用者が端末を過剰に意識してしまい快適でより自然な通信環境を得ることができないという問題もある。

【0007】本発明は、このような問題を解決するためになされたもので、使用者が表示手段の視認可能範囲やビデオカメラの撮像可能範囲を特に意識することなく端末を自然に使用することができる画像・音声通信端末装置を提供することを目的とするものである。

【0008】

【課題を解決するための手段】前記目的を達成するため、本発明では、画像・音声通信端末装置を次の(1)、(2)、(3)のとおり構成する。

【0009】(1)表示方向を変更可能な画像等の表示手段を備えた画像・音声通信端末装置であって、この画像・音声通信端末装置の周囲の複数位置において音声サンプルを拾集する音声サンプル拾集手段と、この音声サンプル拾集手段で拾集した音声サンプルの音量を比較して目標方向を判定する目標方向判定手段と、この目標方向判定手段で判定した目標方向へ前記表示手段の表示方向を変更させる表示方向変更手段とを備えたことを特徴とする画像・音声通信端末装置。

【0010】(2)撮像方向を変更可能なビデオカメラを備えた画像・音声通信端末装置であって、この画像・音声通信端末装置の周囲の複数位置において音声サンプルを拾集する音声サンプル拾集手段と、この音声サンプル拾集手段で拾集した音声サンプルの音量を比較して目標方向を判定する目標方向判定手段と、この目標方向判定手段で判定した目標方向へ前記ビデオカメラの撮像方向を変更させる撮像方向変更手段とを備えたことを特徴とする画像・音声通信端末装置。

【0011】(3)表示方向を変更可能な画像等の表示手段と、撮像方向を変更可能なビデオカメラを備えた画像・音声通信端末装置であって、この画像・音声通信端末装置の周囲の複数位置において音声サンプルを拾集する音声サンプル拾集手段と、この音声サンプル拾集手段で拾集した音声サンプルの音量を比較して目標方向を判定する目標方向判定手段と、この目標方向判定手段で判定した目標方向へ、前記表示手段の表示方向と前記ビデオカメラの撮像方向を変更させる方向変更手段とを備えたことを特徴とする画像・音声通信端末装置。

【0012】

【作用】前記(1)の構成によれば、端末周囲の複数位置で拾集した音声サンプルの音量にもとづいて使用者の方向を判定し、その方向に表示手段の表示方向を向ける。前記(2)の構成によれば、端末周囲の複数位置で

拾集した音声サンプルの音量にもとづいて使用者の方向を判定し、その方向にビデオカメラの撮像方向を向ける。前記(3)の構成によれば、(1)、(2)と同じ判定出力により表示手段とビデオカメラを使用者の方向に向ける。

【0013】

【実施例】以下本発明を実施例により詳しく説明する。

【0014】(実施例1)図1は、実施例1である“画像・音声通信端末装置”の外観を示す斜視図である。同図において、1、2はマイクロホン、8は可動式表示部である。

【0015】図2は、本実施例の構成を示すブロック図である。同図において、1、2は音声入力手段の一つであるマイクロホン、3は音声出力手段の一つであるスピーカである。

【0016】4は、システム制御部14の指示により、音声入出力手段としてマイクロホン1、2、スピーカ3を切り換える切り換え処理、音声入出力手段としてマイクロホン1、2とスピーカ3を使用したときにエコーを消去するためのエコーキャンセル処理、音声サンプル拾集手段としてマイクロホン1、2から拾集した音声データをシステム制御部14に転送する処理、ダイヤルトーン、呼出音、ビジー音、着信音等のトーンの生成処理等を行う音声インタフェース部である。

【0017】5は、システム制御部14の指示により、64kbps PCM A-law、64kbps PCM μ -law、64bps / 56kbps / 48kbps SB-ADPCM、32kbps ADPCM、16kbps (例えばAPC-AB)、8kbps等の音声信号符号化、復号化アルゴリズムに従って、送信音声信号を符号化、受信音声信号を復号化する音声符号化復号化部である。

【0018】6は画像入力手段の一つであり自画像等を入力するためのビデオカメラ、7は画像入力手段の一つであり絵、図面等を入力するための書画カメラ、8は、カメラ6または書画カメラ7よりの入力画像、相手からの受信画像、操作画面等を表示する可動式表示部である。

【0019】9は、システム制御部14の指示により、画像入力手段の切り換え処理、入力画像と受信画像と操作画面の表示切り換え処理及びそれらを表示部8上で分割表示するための画像信号合成処理、表示部8の表示方向を制御するカメラコントロール処理等を行うビデオインタフェース部である。

【0020】10はCCITT勧告草案H.261に従って送信画像の符号化及び受信画像信号の復号化を行うビデオ符号化復号化部、11はデータ通信を行うためのデータ端末である。12は、データ端末11及びシステム制御部14からの送信データを多重分離化部15へ通知するとともに、受信データをデータ端末11またはシステム制御部14へ通知するデータインタフェース部で

ある。

【0021】13は、本装置の制御全般を行うための制御情報入力に使用するキーボード、タッチパネル等の操作部である。14は、CPU、ROM、RAM、補助記憶装置等を備え、各部の状態を監視し装置全体の制御状態に応じた操作/表示画面の作成及びアプリケーションプログラムの実行、音声インタフェース部4への音声サンプル拾集の指示、音声インタフェース部4から転送されてきた音声サンプルの音量データを比較し、その結果により決定された目標の方向をビデオインタフェース部9に指示する等の処理を行うシステム制御部である。

【0022】15は、CCITT勧告H.221に従って、音声符号化復号化部5からの音声信号、ビデオ符号化復号化部10からの画像信号、データインタフェース部12からのデータ、システム制御部14からのBASを送信フレーム単位に多重化するとともに、受信フレームを構成単位の各メディアに分離し各部に通知する多重分離化部である。16は、ISDNユーザ・網インタフェースに従って回線を制御する回線インタフェース部である。

【0023】図3は、本実施例の動作を表すフローチャートである。同図において、S2～S6までの動作を合わせて“表示部自動追尾動作”と呼ぶ。表示部自動追尾動作は、通信開始時に自動的に起動され、また操作部13からの指示で、通信中を含めて随時起動することができる。

【0024】まず回線インタフェース部16が、発信あるいは着信の制御を行い、通信が開始する(S1)。するとシステム制御部14が、音声インタフェース部4を通してマイクロホン1、2に音声サンプル拾集の指示を出し、マイクロホン1、2は音声サンプルの拾集を開始する(S2)。拾集された音声サンプルは音声インタフェース部4を通してシステム制御部14に転送され、システム制御部14はマイクロホン1、2からそれぞれ送られてきた音声サンプルを比較し、どちらの音量が大きいか判定する(S3)。そしてその結果によって、システム制御部14は、話者すなわち目標の方向を決定し(S4)、その方向へ可動式表示部8を向けるようにビデオインタフェース部9へ指示を出す(S5)。ビデオインタフェース部9は指示された方向へ可動式表示部8を向ける(S6)。

【0025】以後、通信終了まで(S8、YES)、操作部13から表示部自動追尾動作の指示があれば(S7、YES)、表示部自動追尾動作S2～S6を繰り返す。

【0026】(実施例2)前記実施例1では、音声サンプル拾集手段として2つのマイクロホンをを用いた。

【0027】これに対し本実施例では、音声サンプル拾集手段として、可動式のマイクロホンを1つ用い、音声サンプル拾集時には、システム制御部14が可動式マイ

クロホンを移動制御することにより、数か所で音声サンプルを拾集する。

【0028】(実施例3)図4は実施例3である“画像・音声通信端末装置”の外観斜視図である。同図において、1、2はマイクロホン、6-1は可動式ビデオカメラである。

【0029】図5は本実施例の構成を示すブロック図である。同図において、6-1、8-1、9-1、14-1以外のブロックは、図1の同符号のブロックと同機能のブロックであり、ここでの説明は省略する。

【0030】6-1は画像入力手段の一つであり自画像等を入力するための可動式ビデオカメラであり、8-1は、カメラ6-1または書画カメラ7よりの入力画像、相手からの受信画像、操作画面等を表示する表示部である。

【0031】9-1は、システム制御部14-1の指示により、画像入力手段の切り換え処理、入力画像と受信画像と操作画面の表示切り換え処理及びそれらを表示部8-1上で分割表示するための画像信号合成処理、可動式ビデオカメラ6の方向を制御するカメラコントロール処理等を行うビデオインタフェース部である。

【0032】14-1は、CPU、ROM、RAM、補助記憶装置等を備え、各部の状態を監視し装置全体の制御状態に応じた操作/表示画面の作成及びアプリケーションプログラムの実行、音声インタフェース部4への音声サンプル拾集の指示、音声インタフェース部4から転送されてきた音声サンプルの音声データを比較し、その結果により決定された目標の方向をビデオインタフェース部9-1に指示する等の処理を行うシステム制御部である。

【0033】図3は、本実施例の動作を表すフローチャートである。同図において、S12～S16までの動作を合わせて“カメラ自動追尾動作”と呼ぶ。カメラ自動追尾動作は通信開始時に自動的に起動され、また操作部13からの指示で、通信中を含めて随時起動することができる。

【0034】まず回線インタフェース部16が、発信あるいは着信の制御を行い、通信が開始する(S11)。するとシステム制御部14-1が、音声インタフェース部4を通してマイクロホン1、2に音声サンプル拾集の指示を出し、マイクロホン1、2は端末周囲の音声サンプルの拾集を開始する(S12)。拾集された音声サンプルは音声インタフェース部4を通してシステム制御部14-1に転送され、システム制御部14-1はマイク

ロホン1、2からそれぞれ送られてきた音声サンプルを比較し、どちらの音量が大きいか判定する(S13)。そしてその結果によって、システム制御部14-1は、話者すなわち目標の方向を決定し(S14)、その方向へ可動式ビデオカメラ6-1を向けるようにビデオインタフェース部9-1へ指示を出す(S15)。ビデオインタフェース部9-1は指示された方向へ可動式ビデオカメラ6-1を向ける(S16)。

【0035】以後、通信終了まで(S18、YES)、操作部13からカメラ自動追尾動作の指示があれば(S17、YES)、カメラ自動追尾動作S12～S16を繰り返す。

【0036】(実施例4)前記実施例3では、音声サンプル拾集手段として2つのマイクロホンをを用いた。

【0037】これに対し本実施例では、この音声サンプル拾集手段として、可動式のマイクロホンを1つ用い、音声サンプル拾集時には、システム制御部14-1がこの可動式マイクロホンを移動制御することにより、数か所で音声サンプルを拾集する。

【0038】なお以上の各実施例は、表示部、ビデオカメラのいずれかを使用者の方向に向けるものであるが、本発明はこれに限らず、表示部、ビデオカメラ共に使用者の方向に向ける形で実施することもできる。また方向変更は取付部を回転する他、ミラーを利用しその反射方向を変更するといった形で実施することもできる。

【0039】

【発明の効果】以上説明したように、本発明によれば、画像等の表示手段、ビデオカメラが使用者を自動追尾するので、使用者は表示手段の視認可能範囲またはビデオカメラの撮像可能範囲を意識しなくてすみ、快適な通信環境が得られる。

【図面の簡単な説明】

【図1】 実施例1の外観斜視図

【図2】 実施例1のブロック図

【図3】 実施例1のフローチャート

【図4】 実施例3の外観斜視図

【図5】 実施例3のブロック図

【図6】 実施例3のフローチャート

【符号の説明】

1、2 マイクロホン

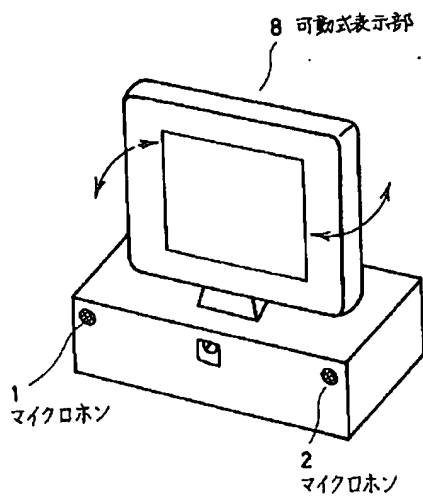
8 可動表示部

9 ビデオインタフェース部

14 システム制御部

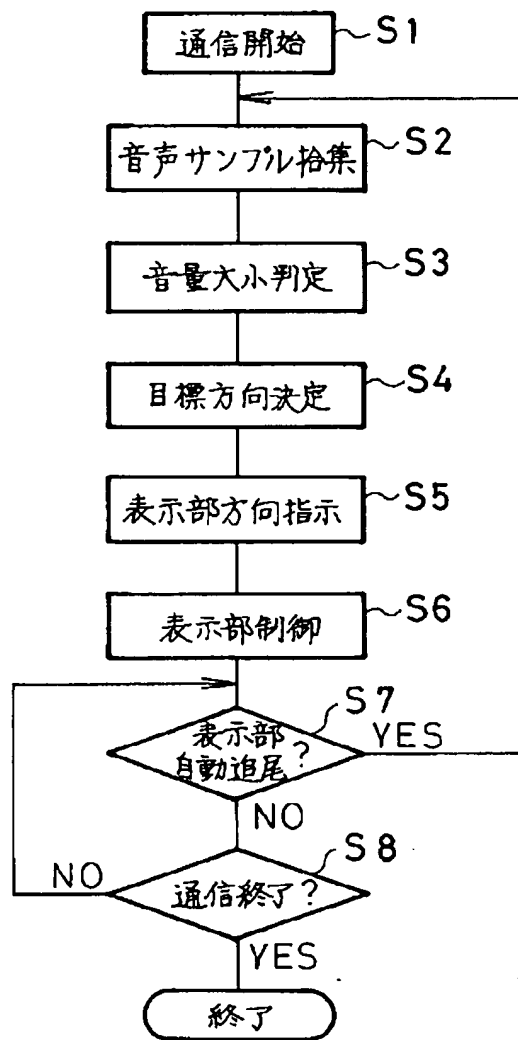
【図1】

実施例1の外観斜視図

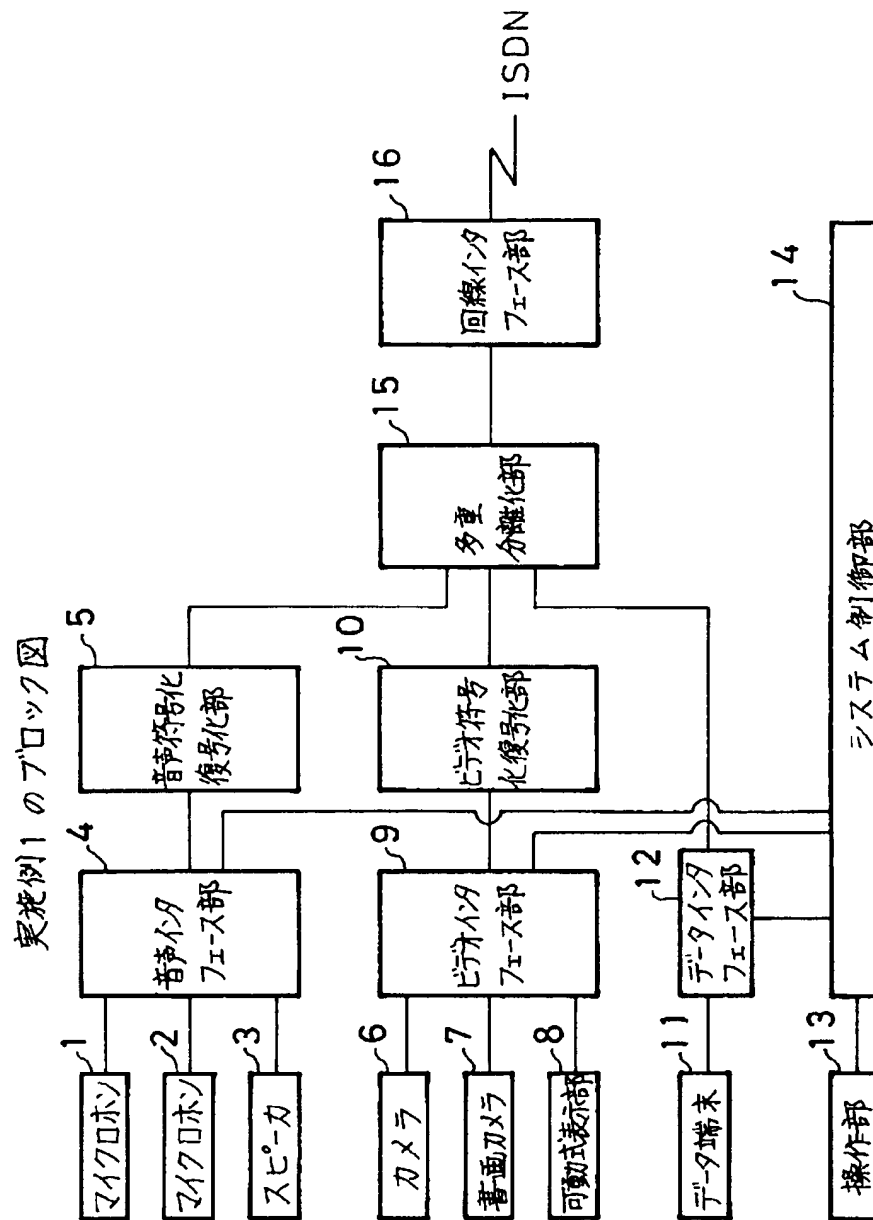


【図3】

実施例1のフローチャート

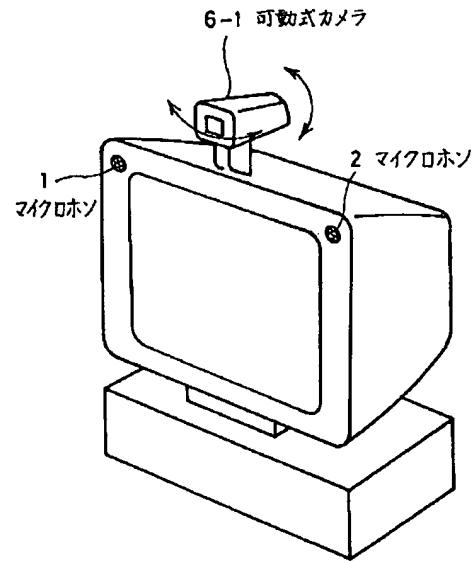


【図2】

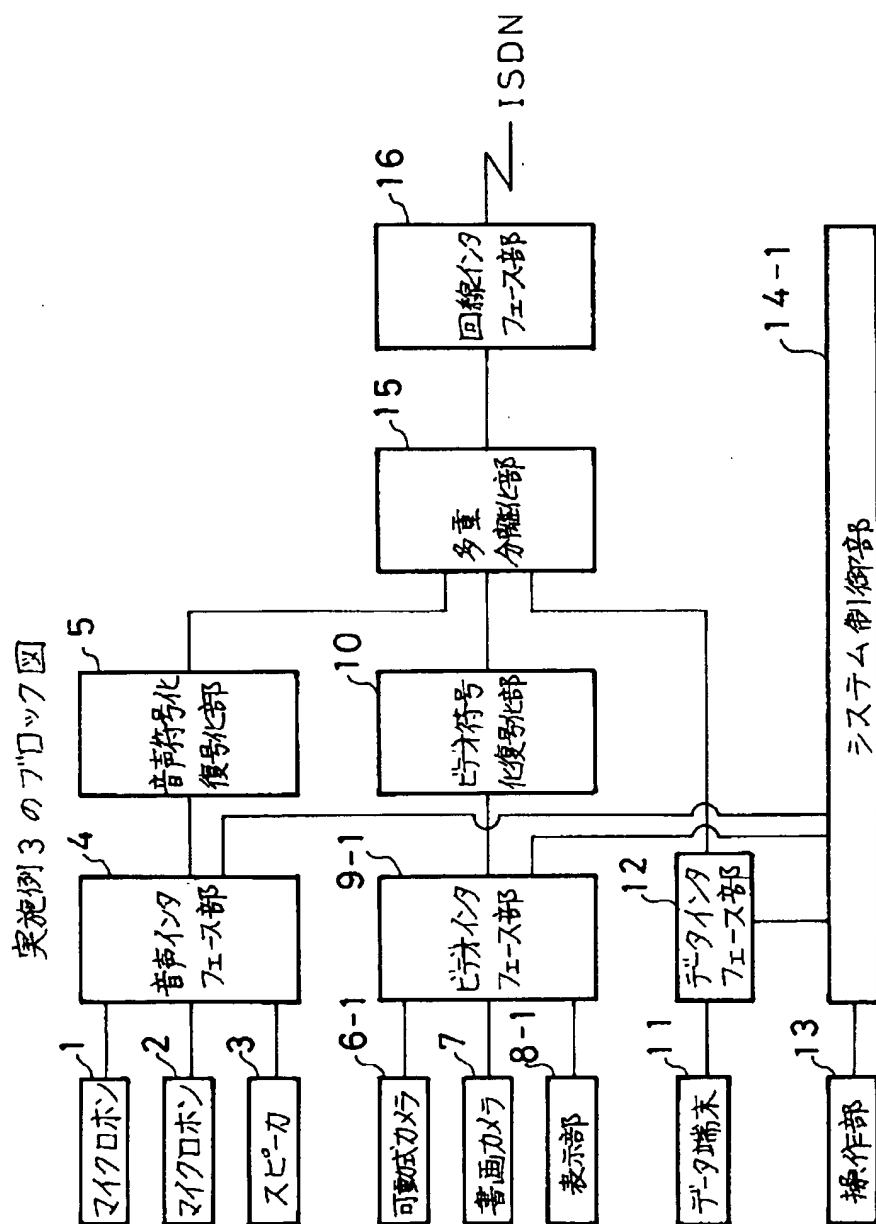


【図4】

実施例3の外観斜視図



【図5】



【図6】

実施例3のフローチャート

